

CODIFICACION ADAPTATIVA DE IMAGENES MEDIANTE LA TRANSFORMADA COSENO DISCRETA*

Luis Torres Urgell Javier Vidal Plana

Dept. Teoría de la Señal y Comunicaciones

ETSIT - UPC

Apdo. 30002 08080 Barcelona

1.- INTRODUCCION

La transmisión digital de imagen juega un papel muy importante en los nuevos sistemas de comunicación visual que se están desarrollando y que en algunos casos ya están en fase de aplicación. Las futuras redes de comunicaciones digitales incluirán, entre otros, comunicaciones de vídeo persona a persona, gráficos, servicios de alta definición, videoconferencia, teletexto y videotexto. Todas estas aplicaciones requieren codificación de imagen con una mayor o menor reducción de la cantidad de bits a transmitir dependiendo de la aplicación en concreto.

De entre las técnicas más conocidas en compresión de imagen destacan las técnicas de transformación (1). La codificación por métodos transformados se basa en dividir la imagen en subimágenes y transformar el conjunto de pixels mediante una transformación lineal, de forma que los nuevos coeficientes sean estadísticamente independientes entre sí. En el receptor se decodifican los coeficientes y se aplica la transformación inversa.

La transformada óptima en el error cuadrático medio es la transformada de Karhunen-Loeve, pero exige un tiempo de cálculo muy elevado. La transformada Coseno Discreta es subóptima pero ofrece un comportamiento excelente tanto en cálculo como en compresión. Los métodos transformados se pueden mejorar con técnicas adaptativas, es decir, dependiendo de la subimagen que se está codificando, se seleccionan más o menos coeficientes en función de una cierta medida de "actividad". Esta medida de actividad indica el grado de variación de los pixels en una determinada subimagen.

En este artículo presentamos una técnica de codificación adaptativa de imágenes fijas utilizando la transformada Coseno Discreta. Como primer paso se divide la imagen en subimágenes cuadradas del mismo número de pixels. A continuación se define una medida de actividad en cada una de dichas subimágenes volviéndose a dividir cada subimagen si dicha medida de actividad excede un determinado umbral. El proceso se aplica tantas veces como sea necesario hasta que se alcanzan subimágenes sin actividad aparente. De esta forma se obtiene un proceso adaptativo e iterativo, que al depender de la imagen, consigue compresiones mayores que las técnicas no adaptativas convencionales. El método es ampliable a secuencias de imagen.

* Este trabajo ha sido subvencionado mediante la ayuda PRONTIC 105/88

2.- METODOS TRANSFORMADOS ADAPTATIVOS

Los métodos adaptativos intentan aprovechar la diferente estructura de las imágenes, para codificarlas de diferente forma según sus características, y de esta forma conseguir una utilización más eficiente de los bits. El objetivo es poner bits sólo donde sean necesarios.

Las imágenes más usuales suelen estar formadas por diferentes regiones dentro de las cuales la variación de la intensidad de los pixels es muy lenta, y separadas unas de otras por contornos frecuentemente bien definidos. Dado que las regiones de poca actividad pueden codificarse utilizando muy pocos bits/pixel, si estas zonas se pueden distinguir de las zonas con contornos, se podrán asignar los bits de una forma efectiva.

Una forma de hacer esto es dividir la imagen en bloques de un tamaño fijo, p.e. 8×8 , y mediante alguna medida de actividad clasificarlos en varios tipos según la actividad que contengan. En este método tiene una gran importancia el tamaño de bloque. Tanto los bloques grandes como pequeños tienen sus ventajas e inconvenientes.

En la figura 1 aparece un dibujo que representa una hipotética imagen formada por dos regiones de intensidad constante, separadas por un contorno abrupto. La imagen se supone 64×64 pixels.

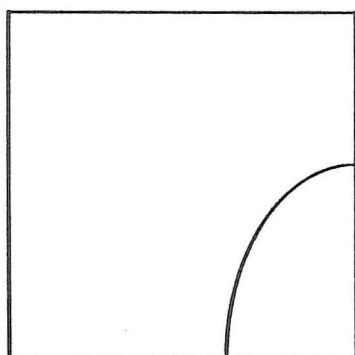


Fig. 1 Contorno en región de intensidad constante.

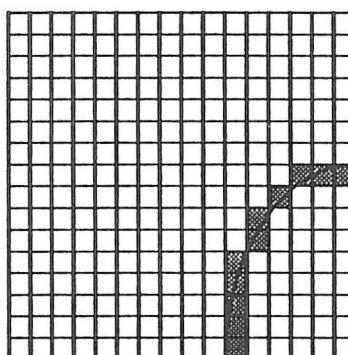


Fig. 2 Clasificación de la imagen en bloques con y sin actividad.

Si ahora se divide la imagen en bloques 4×4 y se clasifican los bloques en dos tipos, bloques con actividad y bloques sin actividad, se obtiene una clasificación como en la figura 2. Los bloques en negrita serían bloques con actividad, el resto bloques sin actividad.

Cómo puede verse, con un bloque muy pequeño se obtiene una adaptación muy eficiente al localizar muy bien el contorno. Una parte muy pequeña de la imagen es clasificada como zona de actividad. El problema de los bloques pequeños es que si no existe actividad no se pueden conseguir grandes compresiones. Por ejemplo, con un bloque 2×2 la mayor compresión posible se obtendría reteniendo un sólo coeficiente y cuantificándolo con 5 bits, ya que con menos bits no queda aceptablemente cuantificado. Aún así el bloque estaría codificado a 1.25 b/p y la calidad sería bastante mala, ya que los bloques siempre contienen algo de actividad y al retener sólo el coeficiente de la componente continua se nota fácilmente el efecto bloque. En general las zonas con poca actividad se pueden comprimir más a medida que tenemos bloques más grandes.

Con un bloque 16×16 la adaptación es mucho menos eficiente. En la figura 3 aparece la clasificación de bloques que se obtiene con un tamaño de 16×16 . Como se

puede ver, existe una proporción mucho mayor de la imagen clasificada como de actividad, cuando realmente el contorno ocupa una zona menor.

El método que aquí se analiza intenta conseguir las ventajas de ambos tamaños de bloque. El método de codificación sería el siguiente : se divide la imagen en bloques muy grandes, por ejemplo 64x64 pixels y se comprueba si cada uno de estos bloques contiene actividad. Si no contiene actividad se codifica el bloque de 64x64 con muy pocos coeficientes. Si contiene actividad, se divide en cuatro bloques de 32x32 pixels y se repite el proceso con cada uno de ellos. Si llegamos a tener bloques 4x4, éstos se codifican con suficientes b/p para que los contornos queden bien codificados. La imagen anterior queda dividida según se indica en la figura 4. Nótese que los bloques sin actividad que vamos localizando se han de codificar con más b/p a medida que son más pequeños. De esta forma se consigue localizar el contorno en bloques pequeños y al mismo tiempo las zonas sin actividad se codifican con bloques grandes que permiten obtener compresiones mayores.

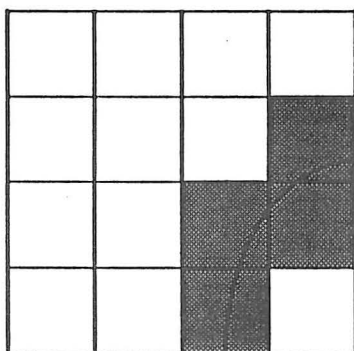


Fig. 3 Clasificación de la imagen en bloques grandes.

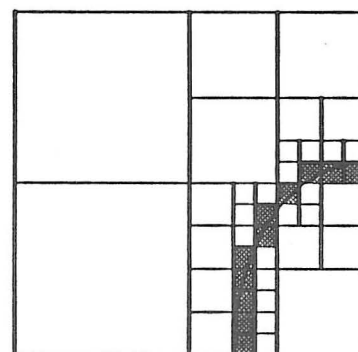


Fig. 4 Clasificación adaptativa de la imagen.

Un elemento muy importante de este método es el detector de actividad. Como detector se ha utilizado el valor medio de los diez valores máximos del resultado de aplicar el algoritmo detector de contornos Sobel (2). El procedimiento es el siguiente:

1) Se calcula el módulo del gradiente en todo el bloque

$$s(x,y) = |f(x+1,y) - f(x,y)| + |f(x,y+1) - f(x,y)|$$

2) Se eligen los diez valores máximos

3) Se halla la media de dichos valores y si ésta supera un umbral, se decide que hay actividad en el bloque.

Si se calculase la media de los valores resultantes del detector Sobel en todo el bloque, en los pixels que presentasen actividad en una pequeña parte del bloque, al estar éste tan promediado, el detector decidiría que no la hay, y la codificación resultante sería de baja calidad. Por ello se usan sólo los diez valores máximos con lo que el detector, es más sensible a pequeñas concentraciones de actividad. Está claro que este valor de diez es arbitrario y se podría haber elegido otro parecido.

3.- RESULTADOS

Para codificar los bloques se han probado dos alternativas; retener un número de coeficientes fijo para cada tamaño de bloque, y retener un número variable para cada tamaño de bloque.

Cuando se utiliza un número fijo, éste se ha elegido de forma experimental, de forma que la actividad que contiene cada bloque quede aceptablemente codificada. Como es práctica habitual en codificación por transformadas, se eligen los coeficientes de mayor varianza. Existe un compromiso entre el número de coeficientes retenido, y el umbral de actividad que se ha fijado para dividir o no un bloque. Si el umbral aumenta se han de retener más coeficientes.

Para cada bloque retenido, se asigna a cada coeficiente un número de bits proporcional al logaritmo de su varianza y se cuantifican con un cuantificador uniforme.

Si el número de dichos coeficientes es variable, se eligen dichos coeficientes de acuerdo con

$$\eta \leq \frac{\sum_{j=1}^Q |y_j|^2}{\sum_{j=1}^N |y_j|^2}$$

para un umbral determinado η . De esta forma se consigue una adaptación adicional para cada bloque.

Cuando se utiliza una asignación variable de coeficientes hay que transmitir también el número de coeficientes que se retienen en cada bloque lo que implica un número de bits adicional.

El algoritmo propuesto requiere transmitir información lateral para especificarle al receptor de que forma va dividida cada imagen. Esta información varía según la imagen vaya más o menos dividida. El caso más desfavorable es cuando la imagen va completamente dividida en bloques 4x4, en cuyo caso es la siguiente : 1 bits para cada bloque 64x64 para especificar si va dividida o no, 1 bits para cada uno de los 4 bloques 32x32 que se generan al dividirlo, 1 bita para cada uno de los 16 bloques 16x16 que se generan al dividir éstos, y así sucesivamente. Los bloques 4x4 no llevan información lateral ya que nunca van divididos. Los bits/píxel de información lateral en el peor de los casos es

$$\frac{1 + 4 + 16 + 64}{4096} = 0.02 \text{ bits/píxel}$$

Por lo tanto esta información no representa una gran sobrecarga adicional de bits.

A continuación pueden verse los resultados que se obtienen al codificar de esta forma la imagen Cheryl. El umbral de actividad se ha seleccionado de 250. Valores similares de umbral proporcionan resultados parecidos, lo que demuestra que la elección del umbral no es un valor crítico siempre y cuando se elija suficientemente alto. La imagen se ha codificado con un número de coeficientes fijo. Empleando un número fijo de coeficientes se obtienen compresiones de 0.3 - 0.4 bits/píxel, mientras que si se utiliza una asignación variable de coeficientes se obtienen compresiones de una calidad algo mejor, con compresiones típicas de 0.5-0.8 bits/píxel. La calidad visual es muy aceptable en todos los casos.



Fig. 5 Imagen original Cheryl



Fig. 6 Cheryl a 0.29 bits/pixel

4.- CONCLUSIONES

Se ha presentado un algoritmo adaptativo utilizando la Transformada Coseno Discreta. La adaptatividad se define a través de un algoritmo detector de contornos clásico como es el de Sobel. La implementación del método es sencilla y se consiguen compresiones del orden de 0.3 bits/pixel con una buena calidad de imagen.

5.- REFERENCIAS

1. P.A. Wintz, " Transform Picture Coding, " Proc. IEEE, vol 60, pp. 809-820, July 1972.
2. R. González, P. Wintz, " Digital Image Processing, " Addison Wesley May 1987.